

## ENCODER/DECODER APPARATUS

Patent number: WO8912292

Publication date: 1989-12-14

Inventor: UNAGAMI SHIGEYUKI (JP); AMANO FUMIO (JP);  
 ISEDA KOHEI (JP); OHTA YASUJI (JP); OKAZAKI KOJI  
 (JP); TANAKA YOSHINORI (JP); TANIGUCHI  
 TOMOHIKO (JP)

Applicant: FUJITSU LTD (JP)

Classification:

- international: G10L9/14

- european: G10L19/04

Application number: WO1989JP00580 19890607

Priority number(s): JP19890061533 19890314; JP19880141343 19880608

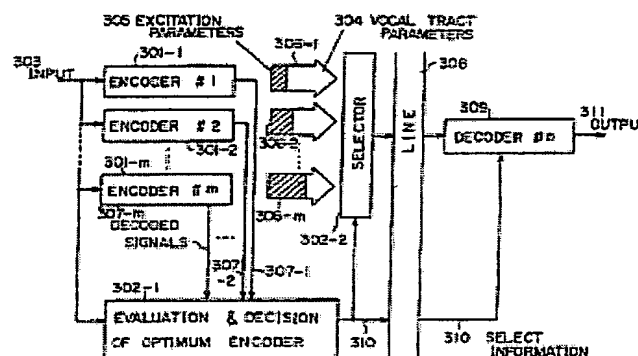
Also published as:

EP0379587 (A)  
 US5115469 (A)  
 EP0379587 (B)

Abstract not available for WO8912292

Abstract of correspondent: **US5115469**

PCT No. PCT/JP89/00580 Sec. 371 Date Feb. 8, 1990 Sec. 102(e) Date Feb. 8, 1990 PCT Filed Jun. 7, 1989 PCT Pub. No. WO89/12292 PCT Pub. Date Dec. 14, 1989. Several encoders perform a local decoding of a speech signal and extract excitation information and vocal tract information from a speech signal for an encoding operation. The transmission rate ratio between the excitation information and the vocal tract information are different for each encoder. An evaluation/selection unit evaluates the quality of decoded signals subjected to a local decoding in each of the encoders, determines the most suitable encoders from among the several encoders based on the result of the evaluation, and selects the most suitable encoder, thereby outputting the selection result as selection information. The decoder decodes a speech signal based on selection information, vocal tract information and excitation information. The evaluation/selection unit selects the output from the encoder in which the quality of a locally decoded signal is the most preferable. When vocal tract information changes little, the vocal tract information is not output, thereby allowing for increased quality of information. As much of the surplus of unused vocal tract information as possible is assigned to a residual signal. Thus, the quality of a decoded speech signal is improved.



Data supplied from the esp@cenet database - Worldwide

BEST AVAILABLE COPY

## 公表特許公報(A)

平2-502491

④公表 平成2年(1990)8月9日

⑤Int. Cl.<sup>9</sup>G 10 L 9/14  
9/18

識別記号

E  
C

庁内整理番号

8622-5D  
8622-5D

審査請求有

予備審査請求 未請求

部門(区分) 6(2)

(全 11 頁)

④発明の名称 符号化/復号化装置

②特 願 平1-506723

⑤②出 願 平1(1989)6月7日

⑤翻訳文提出日 平2(1990)2月1日

⑥国際出願 PCT/JP89/00580

⑦国際公開番号 WO89/12292

⑦国際公開日 平1(1989)12月14日

優先権主張 ②昭63(1988)6月8日③日本(J P)④特願 昭63-141343

⑦発明者 谷 口 智 彦

神奈川県横浜市港北区日吉2-18-22 メゾンド田園

⑦発明者 伊 勢 田 衡 平

神奈川県川崎市中原区井田770 コーボウイスタリア

⑦発明者 岡 崎 晃 二

神奈川県川崎市幸区柳町18 第一両角荘102

⑦発明者 天 野 文 雄

東京都世田谷区千歳台1-17-1

⑦発明者 海 上 重 之

神奈川県厚木市森の里2-34-10

⑦出願人 富士通株式会社

神奈川県川崎市中原区小田中1015番地

⑦代理人 弁理士 井 桁 貞一

⑧指定国 AT(広域特許), BE(広域特許), CH(広域特許), DE(広域特許), FR(広域特許), GB(広域特許), I T  
(広域特許), J P, L U(広域特許), NL(広域特許), SE(広域特許), U S

最終頁に続く

## 請 求 の 範 囲

1. 音声信号について、該音声信号の特性を前記音声信号の調音特性を表す調音情報と前記音声信号の音源特性を表す音源情報に分離することにより、符号化を行う音声符号化装置であって、

各々、前記音声信号の局部復号化を行いながら、前記音声信号から調音情報及び音源情報を抽出して符号化し、該符号化された各情報間の伝送速度比が各々異なる複数の符号化手段と、

該各符号化手段で局部復号化が行われた各復号化信号の品質を評価し、該評価結果に基づいて前記符号化手段のうち最適な符号化手段を決定・選択すると共に、該選択結果を選択情報として出力する評価・選択手段と、

を含み、

該評価・選択手段で選択された符号化手段で符号化された前記調音情報と音源情報、並びに前記評価・選択手段からの選択情報を出力する、

ことを特徴とする音声符号化装置。

2. 請求項1記載の装置であって、

音声信号について、該音声信号の特性を前記音声信号の声道特性を表すL P Cパラメータと前記音声信号の音源特性を表す残差信号に分離することにより符号化を行う、

ことを特徴とする音声符号化装置。

3. 請求項1記載の装置であって、

前記評価・選択手段は、前記音声信号に対する前記各復号化信号の波形歪を計算することにより前記各復号化信号の品質を評価し、

該波形歪が小さい方の復号化信号に対応する前記符号化手段を決定・選択する、

ことを特徴とする音声符号化装置。

4. 請求項1記載の装置であって、

前記評価・選択手段は、前記音声信号に対する前記各復号化信号のスペクトル歪を計算することにより前記各復号化信号の品質を評価し、

該スペクトル歪が小さい方の復号化信号に対応する前記部分符号化手段を決定・選択する、

ことを特徴とする音声符号化装置。

5. 請求項1記載の装置であって、

前記評価・選択手段は、前記音声信号に対する前記各復号化信号の波形歪及びスペクトル歪を計算することにより前記各復号化信号の品質を評価し、

該波形歪及びスペクトル歪に基づいて前記符号化手段を決定・選択する、

ことを特徴とする音声符号化装置。

6. 音声信号について、所定フレーム毎に前記音声信号の特性を前記音声信号の声道特性を表すL P Cパラメータと前記音声信号の音源特性を表す残差信号に分離することにより、符号化を行う音声符号化装置であって、

前記音声信号の局部復号化を行いながら、前記各フレ

ム毎に前記音声信号からLPCパラメータ及び残差信号を抽出して符号化する第1の符号化手段と、

前記音声信号の局部復号化を行いながら、前記第1の符号化手段で求まる現在より以前のフレームのLPCパラメータを用いて前記音声信号から前記残差信号を抽出して符号化する第2の符号化手段と、

前記第1及び第2の符号化手段で局部復号化が行われた各復号化信号の品質を評価し、該評価結果に基づいて前記第1及び第2の符号化手段のうち適する方の符号化手段を決定・選択する評価・選択手段と、

を含み、

該評価・選択手段で前記第1の符号化手段が選択された場合は、該第1の符号化手段で符号化された前記LPCパラメータと残差信号、並びに前記評価・選択手段からの選択情報を出力し、前記評価・選択手段で前記第2の符号化手段が選択された場合は、該第2の符号化手段で符号化された前記残差信号及び前記評価・選択手段からの選択情報を出力する、

ことを特徴とする音声符号化装置。

7. 請求項6記載の装置であって、

前記評価・選択手段は、

前記音声信号に対する前記各復号化信号の波形歪及びスペクトル歪を計算することにより前記各復号化信号の品質を評価し、

前記第1の符号化手段の復号化信号の波形歪の方が小さ

い場合は該第1の符号化手段を決定・選択し、

前記第2の符号化手段の復号化信号の波形歪の方が小さくかつ前記第1の符号化手段の復号化信号のスペクトル歪の方が小さい場合は前記第1の符号化手段を決定・選択し、

前記第2の符号化手段の復号化信号の波形歪の方が小さくかつ前記第2の符号化手段の復号化信号のスペクトル歪の方が小さい場合は前記第2の符号化手段を決定・選択する、

ことを特徴とする音声符号化装置。

8. 音声信号の復号化を行う音声復号化装置であって、

選択状態が第1の内容を示している場合には、現在のフレームの符号化されたLPCパラメータ及び符号化された残差信号を入力して音声信号を復号化する第1の復号化手段と、

選択状態が第2の内容を示している場合に、現在のフレーム以前に入力済みの符号化されたLPCパラメータ及び現在のフレームの符号化された残差信号から音声信号を復号化する第2の部分復号化手段と、

を含むことを特徴とする音声復号化装置。

9. 音声信号について、該音声信号の特性を前記音声信号の調音特性を表す調音情報と前記音声信号の音源特性を表す音源情報に分離することにより、符号化を行い、該音声符号化装置で符号化された前記符号化された調音情報及び符号化された音源情報に基づいて前記音声信号の復号化を行う音声符号化／復号化装置であって、

## 明 細 書

(発明の名称)

符号化／復号化装置

(技術分野)

本発明は、音声信号について情報圧縮処理を施してから伝送を行うための音声符号化方式に関する。

近年、企業内通信システムやデジタル移動無線システムあるいは音声蓄積システムなどにおいては、音声情報に対して4～16Kbps程度の情報量で高効率な圧縮を行う音声符号化方式が要求されている。

(背景技術)

音声の予測符号化方式の第1の従来例として、予測器の予測パラメータ(音道情報)と残差信号(音源情報)を多重化して送信側へ伝送する適応予測符号化方式がある。

第1図は第1の従来例に係る音声符号化方式を示すブロック図であるが、かかる符号化方式に使用される符号化器100は、線形予測分析部101、予測器102、量子化器103、多重化部104、加算器105及び106を有している。

ここで、線形予測分析部101は入力音声信号について分析を行うことにより予測パラメータを出力するもので、予測器102は後述する加算器106の出力と線形予測分

各々、前記音声信号の局部復号化を行いながら、前記音声信号から調音情報及び音源情報を抽出して符号化し、該符号化された各情報間の伝送速度比が各々異なる複数の符号化手段と、

該各符号化手段で局部復号化が行われた各復号化信号の品質を評価し、該評価結果に基づいて前記符号化手段のうち最適な符号化手段を決定・選択すると共に、該選択結果を選択情報として出力する評価・選択手段と、

前記評価・選択手段からの選択情報並びに前記評価・選択手段で選択された符号化手段で符号化された前記調音情報と音源情報から前記音声信号を復号化する復号化手段と、を含むことを特徴とする音声符号化／復号化装置。

析部101からの予測パラメータを用いて入力信号を予測するもので、加算器105は入力信号と予測信号との差を演算して誤差情報を出力し、量子化部103はその誤差情報について量子化を施して残差信号を得るもので、加算器106は予測器102と量子化器103の出力を加算して予測器102にフィードバックするもので、多重化部104は線形予測分析部101からの予測パラメータと量子化器103からの残差信号とを多重化して受信側へ伝送するものである。

このような構成により、入力信号は線形予測分析部101で一定のフレーム周期毎に線形予測分析されることにより、予測パラメータが適当なビットが(図示しない)符号器で割り当てられた声道情報として抽出される。そして、この予測パラメータは、このようにして符号化された後、予測器102及び多重化部104に出力される。予測器102では上記予測パラメータ及び加算器106の出力に基づいて入力信号が予測される。続いて、この予測された情報と入力信号との差である誤差情報が加算器105で演算され、この誤差情報に対して量子化器103で量子化処理が施され、適当なビットが割り当てられて符号化されて残差信号が得られ、この信号が音源情報として多重化部104に出力される。

その後は、上記符号化された予測パラメータ及び残差信号が、多重化部104で多重化されて受信側へ伝送される。

一方、加算器106では、予測器102が予測した入力

02でベクトル量子化されて予測器203に与えられる。また図示しないが入力信号からピッチ周期 $m$ およびピッチ係数 $C$ 、及びゲイン $G$ も抽出される。

また、白色雑音コードブック204から残差波形パターン(コードベクトル)が逐次読み出され、各パターンはまず乗算器205に入力し、ここでゲイン $G$ が乗算される。その出力は、ディレイ回路206、乗算器207、加算器208からなるフィードバックループ、すなわち長期予測部に入力し、ここで残差信号が合成される。なお、ディレイ回路206のディレイ量は前記ピッチ周期と同じ値に設定され、乗算器207ではディレイ206の出力値に前記ピッチ係数 $C$ が乗算される。

続いて、加算器208から出力される合成された残差信号は、予測器203、加算器209からなるフィードバックループ、すなわち短期予測部に入力し、ここで予測された入力信号が合成される。このときの予測パラメータは前記量子化部202からのLPCパラメータである。上記予測入力信号は減算器210において入力信号から差し引かれ、誤差信号が得られる。その誤差信号に対し、重み関数器211で、人間の聴覚特性を考慮した重み付けがされる。これは、人間の耳に対する誤差の影響の仕方が、周波数帯域によって異なるため、その影響の仕方を均一にするための補正処理である。

そして、重み関数器211の出力は、エラー電力評価部212に入力し、ここで各フレーム内でのエラー電力が評

価と量子化器103で量子化された残差信号が加算され、この加算出力が再び予測器102に入力して、前記予測パラメータと共に前記入力信号側に用いられる。

この場合、1フレーム当たりの予測パラメータに対する割当てビット数を $\alpha$ ビット/フレームに固定し、残差信号に対する割当てビット数を $\beta$ ビット/フレームに固定している。従って、送信側に伝送されるのは $(\alpha + \beta)$ ビット/フレームで、この場合、伝送速度は例えば8 Kbpsとなる。

次に、第2図は、第2の従来例に係る音声符号化方式を示すブロック図である。この従来例は、低ビットレート音声符号器の1つとして知られる。CELP(Code Excited Linear Prediction)符号器である。

CELP符号器は、基本的には、第1図に示した前記第1の従来例と同様、LPC方式に基づいて得られるLPCパラメータ(予測パラメータ)と残差信号を符号化、伝送する方式である。ただし、この方式では、残差信号はコードブック内の残差パターンによって高効率符号化されることが特徴である。

CELP符号器の詳細はAtal, B.S. and Schroeder, M.R., "stochastic coding of speech at very low bit rate", Proc. ICASSP 84, 1610~1613, 1984.にあるが、第2図で概略を説明すると次の如くである。

入力信号はLPC分析部201でLPC分析され、分析されたLPCパラメータ(予測パラメータ)は量子化部2

価される。

今、白色雑音コードブック204は、複数の残差波形パターン(コードベクトル)のサンプルを持っている。そして、上記一連の処理は、それらのサンプルの全てについて繰り返され、前記エラー電力が最小となる残差波形パターンがそのフレームの残差波形パターンとして選択される。

以上のようにして、各フレーム毎に求まる残差波形パターンは、前記量子化器202からのLPCパラメータ及び前記ピッチ周期 $m$ 、ピッチ係数 $C$ 、及びゲイン $G$ と共に、受信側に伝送される。一方、受信側については特に図示していないが、受信側では伝送されてきたピッチ周期 $m$ とピッチ係数 $C$ により、前述したのと同様の長期予測部が構成され、それに伝送されてきたインデックスに対応する残差波形パターンが入力されることにより、残差信号が再生される。更に、伝送されてきたLPCパラメータにより前述したのと同様の短期予測部が構成され、それに上記再生された残差信号が入力されることにより、入力信号が再生される。

ところで、人間の発声機構における音源部と声道部との各動特性はそれぞれ異なるため、各時点においてそれぞれが伝送すべき情報量は異なるという性質を有する。

しかし、上記のような第1図又は第2図の従来の音声符号化方式では、音源情報と声道情報とを固定した比率の情報量で伝送しているため、音声についての上記のような性質を利用していないことになる。これにより、伝送速度を

低くすると、量子化が粗くなって、雑音が増え、良好な音声品質を維持することが困難になるという問題点がある。

上記問題点と、第1図又は第2図の従来例について説明すると次のようになる。

すなわち、音声信号には、その性質が急激に変化する期間と準定状态的な区間が存在し、後者では予測パラメータの値は余り変わらない。即ち予測パラメータ(LPCパラメータ)もフレーム間の相関が余りない場合と相関が強い場合があるが、従来では各フレーム毎に一定の情報量で予測パラメータ(LPCパラメータ)を伝送しており、上記の音声信号の性質で余り利用していない。このため伝送情報に冗長性が生じ、伝送情報量の割には、受信側復号器での再生音声の品質が十分でない。

#### (発明の開示)

本発明は、音源情報と声道情報との間の伝送すべき情報量の比率について複数種類のモードを用意し、音声符号化時に、最良な再生音声品質が得られるモードに切り替えられるようにした、モード切替型音声符号化復号化方式を提供することを目的とする。

特に、本発明では伝送情報の冗長性を抑えて、即ち余り変化しない声道情報なら送らないようにし、代わって他の音質向上に役立つ音源情報等に多くのビット数を割り当てて受信側再生音声の品質向上を図ることを目的とする。上記目的を達成するために、本発明は、以下に示す構成を開

差信号を抽出して符号化する第2の部分符号化手段とからなる。

次に、評価・選択手段302-1、302-2は、各符号化手段301-1~301-mで局部復号化が行われた各復号化信号307-1~307-mの品質を評価し、該評価結果に基づいて複数の符号化手段301-1~301-mのうち最適な符号化手段を決定し、選択すると共に、該選択結果を選択情報310として出力する。評価・選択手段は、例えば第3図の如く、評価・決定部302-1と選択部302-2からなる。

上記構成の音声符号化装置は、評価・選択手段302-1、302-2で選択された符号化手段で符号化された調音情報304と音源情報305、並びに評価・選択手段302-1、302-2からの選択情報301、例えば回線308に出力するものである。

復号化手段309は、上述の音声符号化装置から伝送されてきた選択情報310、調音情報304及び音源情報305から音声信号311を復号化する。

上述の構成により、評価・選択手段302-1、302-2は、局部復号化された復号化信号307-1~307-mによって品質が良いと評価された符号化手段の符号化出力304及び305を選択する。

従って、音声信号の調音特性にあまり変化のない部分では、LPCパラメータを出力せず、これにより生じた情報量の余裕分を残差信号の方に、より多く割り当てることが

示する。

まず、本発明は、音声信号について、該音声信号の特性を前記音声信号の調音特性を表す調音情報(一般的に、声道情報)と音声信号の音源特性を表す音源情報に分離することにより、符号化を行う音声符号化装置を前提とする。ここで、調音特性は、人間の声道及び鼻等によって形成される音声の周波数特性であり、近似的には声道特性のみを指す場合が多い。そして、その声道特性を表す声道情報は例えば音声信号を線形予測分析して得られるLPCパラメータである。また、音源情報は、例えば残差信号である。上記音声符号化装置に対応する音声復号化装置も前提とする。このような音声符号化/復号化装置を前提として、本発明は第3図に示す構成を有する。

複数の符号化手段301-1~301-mは、各々、音声信号303の局部復号化を行いながら、音声信号303から調音情報(声道情報)304及び音源情報305を抽出して符号化する。これらは、一般的にはパラメータ化される場合が多い。そして、各々で符号化された各情報間の伝送速度比は、第3図の306-1~306-mで概念的に示されるように、各々異なっている。符号化手段は、例えば音声信号の局部復号化を行いながら、各フレーム毎に音声信号からLPCパラメータ及び残差信号を抽出して符号化する第1の部分符号化手段と、音声信号の局部復号化を行いながら、第1の部分符号化手段で求まる現在より以前のフレームのLPCパラメータを用いて音声信号から残

できる。これにより、音声復号化装置において復号化される復号化音声311の品質を改善することが可能となる。

第3図に示す原理ブロック図では、音声符号化装置と音声復号化装置が回線308を介して超て構成されているが、音声符号化装置のみ、又は音声復号化装置のみで構成することも可能である。この場合、音声符号化装置の出力は例えばメモリに蓄積され、音声復号化装置への入力例えばメモリからの出力である。

声道情報は、線形予測分析に基づくLPCパラメータに限られるものではなく、例えばケプストラム分析に基づくケプストラムパラメータ等であってもよい。また、残差信号の符号化方式としては、残差信号をピッチ情報と雑音情報とに分離して符号化する方式、CEL P符号化方式のほか、REL P(Residual Excited Linear Prediction)符号化方式等が適用できる。

#### (図面の簡単な説明)

第1図は、第1の従来例を示すブロック図、

第2図は、第2の従来例を示すブロック図、

第3図は、本発明の原理ブロック図、

第4図は、第1の実施例を示すブロック図、

第5図は、第2の実施例を示すブロック図、

第6図は、第2の実施例の動作フローチャート、

第7A図は、第2の従来例の伝送ビット割り当てを示した図、

第7B図は、第2の実施例の伝送ビット割り当てを示した図である。

(発明を実施するための最良の形態)

以下、図面を用いて本発明の実施例を説明する。

第4図は本発明の第1の実施例の構成図であり、本実施例は第1図に示す第1の従来例に対応する。

第4図の上段の量子化器403-1、予測器404-1、加算器405-1、406-1、LPC分析部402は第1図の103、102、105、106、101と同様であり、適応予測音声符号器を構成する。本実施例ではこれに更に第2の量子化器403-2、予測器404-2、加算器405-2、406-2を設ける。予測器404-2に与えるLPCパラメータは、LPC分析部402で得たものをスイッチ411の端子Aを介してフレーム遅延回路411で遅延させたものとする。第4図上段の第1図と同じ部分が出力端408、409から送出するのは、LPCパラメータと残差信号であり、これをAモードとする。第4図下段の本実施例で付加した部分が出力端412から送出するのは残差信号のみであり、これをBモードとする。評価部407-1、407-2はA、Bモードの符号器のS/Nを評価し、モード決定部413は、その評価結果により相手側(受信側)へ送出するためのA、Bモードいずれにするかを決定する信号A/Bを生じる。スイッチ(SW)部410は、前フレームでAモードが選ばれた場合A

側を選択し、現フレームのBモードのLPCパラメータとして1フレーム前のAモードの値を使用する。また、前フレームでBモードが選ばれた場合はB側を選択し、1フレーム前のBモードの値、すなわち数フレーム前のAモードの値を使用するように働く。

この回路構成では、毎フレーム、A、B両モードの符号器が並列に動作する。Aモード符号器では、声道情報等として現在のフレームの予測パラメータ(LPCパラメータ)を出力端409から出力し、音源情報として残差信号を出力端408から出力する。この場合の伝送速度はLPCパラメータが $\beta$ ビット/フレーム、残差信号が $\alpha$ ビット/フレームとする。Bモード符号器では1フレーム又は数フレーム前のLPCパラメータを用いて得た残差信号を出力端412から出力する。この場合の残差信号の伝送速度は $(\alpha + \beta)$ ビット/フレームとする。つまり、余り変わらないので送らないLPCパラメータのビット数だけ残差信号のビット数を多くする。予測器404-1、404-2への入力信号は加算器406-1、406-2の出力端からの局部復号出力であって、受信側で復号される信号に等しい。そして、評価部407-1、407-2はこの局部復号出力と入力端401の入力信号を比較して復号音声の品質を評価する。この評価には、例えばフレーム内での信号対量子化雑音比(SNR)が用いられ、評価部407-1、407-2からSN(A)、SN(B)として出力される。モード決定部413はこれらと比較してSN(A)

$> \text{SN}(B)$  ならAモードを指示する信号を、SN(A)  
 $< \text{SN}(B)$  ならBモードを指示する信号を出力する。

上記モード決定部413からのAモード又はBモードを指示する信号は、特に図示しないセレクトクに入力する。セレクトクには、出力端408、409、412の各出力が入力している。そして、このセレクトクは、Aモードが指示されたら、出力端408、409からの残差信号とLPCパラメータを選択して相手側に出力し、Bモードが指示されたら、出力端412からの残差信号を選択して相手側に出力する。

A、Bモードの選択は、フレーム毎に行う。そしていずれのモードであっても伝送速度は前述のように $(\alpha + \beta)$ ビット/フレームであって、変わらない。この $(\alpha + \beta)$ ビット/フレームのデータに、当該データはA、Bいずれのモードかを示す信号A/Bをあらわす1ビット/フレームを $(\alpha + \beta)$ ビット/フレームに付加して受信側へ送出する。

本発明の構成では、Bモードの方が品質良好ならBモードを送る、のであるから、再生音声品質は第1図の第1の従来方式に比べて良くなることはあっても悪くなることはないという利点を得られる。

次に、第5図は本発明の第2の実施例の構成図である。この実施例は、第2図の第2の従来例に対応するものである。同図において、501-1、501-2は符号器で、これらの符号器としては第2図に示したCELP符号器が

使用される。そして、一方の符号器501-1は、音声例えば10~30ms毎に区切ったフレーム毎に線形予測分析等を行い、その分析結果として予測パラメータと、残差波形パターン、ピッチ周波数、ピッチ係数及びゲインとを出力するもので、他方の符号器501-2は、線形予測分析は行わず、残差波形パターン等のみを出力するものである。従って、後述するように符号器501-2の方が残差波形パターンに対して量子化ビット数を符号器501-1よりも多く割り当てることができる。

以下、符号器501-1による動作モードをAモードと呼び、符号器501-2による動作モードをBモードと呼ぶ。

符号器501-1において、線形予測分析部506は、LPC分析部201と量子化部202を合わせた機能を有する。白色雑音コードブック507-1、ゲイン調整器508-1及び誤差演算部511-1は、各々、第2図の204、205及び210に対応する。また、長期予測部509-1は、第2図の206~208に対応し、第2の従来例で前述したようにピッチ情報を入力して音源として機能する。更に、短期予測部510-1は、第2図の203及び209に対応し、第2の従来例で前述したように予測パラメータを入力して声道として機能する。加えて、誤差評価部512-1は、第2図の211及び212に対応し、第2の従来例で前述したようにエラー電力の評価を行う。この場合、誤差評価部512-1は、白色雑音コードブッ

ク507-1'の各アドレス(位相)を順次指定しながら、第2の従来例で前述したように、そのコードベクトル(残差パターン)についてエネルギーの評価を行い、それが最小となるコードベクトルを選択し、その番号すなわち白色雑音コードブック507-1中の残差波形パターンの番号が残差信号情報として出力される。

なお、誤差評価部512-1からはフレーム内の波形歪情報を有するセグメンタルS/N(S/N<sub>a</sub>)も出力されるようになっている。

また、この符号器501-1においては、第2図で前述したように、線形予測分析部506からの予測パラメータ(LPCパラメータ)のほか、特に図示しない分析部からのピッチ周期、ピッチ係数及びゲインも符号化されて出力される。

次に、符号器501-2において、507-2~512-2で示した各部は、符号器501-1の507-1~512-1で示した各部と全く同様である。そして、符号器501-2は、線形予測分析部506を持たず、代わりに係数メモリ513を有する。係数メモリ513は、線形予測分析部506からの予測係数(予測パラメータ)を記憶するもので、この係数メモリ513からの情報が短期予測部510-2へ線形予測パラメータ情報として供給される。

なお、この係数メモリ513は、Aモードが発生するたびに(即ち符号器501-1からの出力が選択されるたびに)、更新され、Bモードのとき(即ち符号器501-2

パラメータから第2のLPCケプストラム係数を求める。そして、上記第1及び第2のLPCケプストラム係数から現在のフレームにおけるLPCケプストラム距離CDを計算する。このようにして求まるLPCケプストラム距離は、上記2組のLPCパラメータによって定まる各声道スペクトル特性間の差すなわちスペクトル歪を良く表現することが、一般に知られている。

動作モード判定部516は、各符号器501-1、501-2からのセグメンタルS/N、S/N<sub>a</sub>とケプストラム計算部516からのLPCケプストラム距離(CD)とを受けて、第6図の動作フローチャートで示される処理を行う。

尚、この動作については後述する。

動作モード判定部516によってAモード(符号器501-1)が選択された場合は、スイッチ514は、Aモード端子側に切り替わり、動作モード判定部516によってBモード(符号器501-2)が選択された場合は、Bモード端子側に切り替わる。そして、このスイッチ514の切替動作によって、係数メモリ513は、Aモードが発生するたびに(即ち符号器501-1からの出力が選択されるたびに)、更新され、Bモードのとき(即ち符号器501-2からの出力が選択されるとき)は、更新されずにその値を保持する。多重化部504(MUX)は、符号器501-1からの残差信号情報と予測パラメータとを多重化するものである。なお、セレクト517は、多重化部504

からの出力が選択されるとき)は、更新されずにその値を保持するようになっており、これにより、係数メモリ513の中には常に復号器518(受信側)へ伝送された最新の予測係数が保持されるようになっている。

上記構成により、符号器501-2は、予測パラメータは出力せず、残差信号情報、ピッチ周期、ピッチ係数及びゲインを出力することになる。従って、後述するように予測パラメータを出力しない分、残差信号情報に多くのビットを割り当てることが可能となる。

品質評価・符号器選択部502は、フレームごとに各符号器501-1、501-2で局部復号化された結果に基づいて、再生音声品質の最良な符号器501-1又は501-2を選択するものである。この品質評価・符号器選択部502においては、各符号器501-1又は501-2における再生音声の品質評価のために、再生音声信号A、Bの波形歪およびスペクトル歪が併用される。即ち、再生音声品質評価のために、各フレーム内のセグメンタルS/NとLPCケプストラム距離(CD)とが併用される。

このために、この品質評価・符号器選択部502は、ケプストラム計算部515、動作モード判定部516、スイッチ514を備えて構成されている。

ここで、ケプストラム計算部515は、線形予測分析部506からの現在のフレームに対応するLPCパラメータから第1のLPCケプストラム係数を求め、また、係数メモリ513からの、Bモードで現在使用されているLPC

からの符号化部501-1についての多重化出力(残差信号情報と予測パラメータとの多重化出力)と、符号器501-2からの残差信号情報出力とを、動作モード判定部516からの符号器番号情報に基づいて選択して出力するものである。

518は復号器で、この復号器518は、伝送路を介して伝送されてくる符号器501-1からの残差信号情報と予測パラメータあるいは符号器501-2からの残差信号情報に基づき、再生音声信号を出力するもので、このために復号器518は、符号器501-1、501-2における白色雑音コードブック507-1、507-2、長期予測部509-1、509-2、短期予測部510-1、510-2等と同等のものを有している。

分離部505(DMUX)は、符号器501-1から送られてくる残差信号情報と予測パラメータとを多重化した信号を残差信号情報と予測パラメータとに分離するものである。

なお、第5図において、伝送路503を挟んで、左側に位置するものが送信側に設けられ、右側に位置するものが受信側に設けられる。

上述の構成により、音声入力信号が、符号器501-1では、予測パラメータと残差信号とについて符号化され、一方、符号器501-2では、残差信号についてだけ符号化される。そして、このとき品質評価・符号器選択部502によって、フレームごとのセグメンタルS/N情報とL

PCケブストラム距離情報とに基づき、再生音声品質の最適な符号器501-1または501-2の符号器番号iが選択される。すなわち、品質評価・符号器選択部502内の動作モード判定部516は、第6図で示される動作フローチャートに基づいて、以下に示す処理を実行する。

符号器501-1または501-2は、符号器番号iを入力することによって選択され、Aモードでは $i=1$ であり、Bモードでは $i=2$ である。

まず、 $S/N_A > S/N_B$ 、すなわち符号器501-1におけるセグメンタル $S/N$ の方が良ければ、Aモードを選択し、符号器501-1を示す符号器番号 $i(=1)$ をセレクト517に出力する(第6図、S1→S2)。

逆に、 $S/N_A < S/N_B$ 、すなわち符号器501-2におけるセグメンタル $S/N$ の方が良ければ、更に、次の判別を行う。すなわち、ケブストラム計算部515からのLPCケブストラム距離CDと所定の閾値 $CD_{TH}$ を比較し(S3)、CDがこの閾値 $CD_{TH}$ よりも小さい場合、すなわちスペクトル歪が小さい場合は、Bモードを選択し、符号器501-2を示す符号器番号 $i(=2)$ をセレクト517に出力する(S4)。CDが上記閾値 $CD_{TH}$ よりも大きい場合、すなわちスペクトル歪が大きい場合は、Aモードを選択し、符号器501-1を示す符号器番号 $i(=1)$ をセレクト516に出力する(S3→S2)。

以上の動作により、最適な符号器が選択される。

このように、2つの評価関数を併用するのは、以下の理

り違っておらず、現在のフレームの予測スペクトルも前のフレームのスペクトルとあまり違っていない。そこで、残差信号情報のみが符号器501-2から伝送されることになる。この場合より多くの量子化ビットが残差信号に割り当てられ、残差信号の量子化の品質はより良くなり、予測パラメータと残差信号の両方が反対側のステーションに伝送される場合よりも、比較的多数のビットが伝送できる。例えば、同じ「あー」という音が一連のフレーム以上で発音され続けると、Bモード(符号器501-2)が有効に使用される。

なお、前述のように符号器501-2の係数メモリ513は、Aモードが発生するたびに(即ち符号器501-1からの出力が選択されるたびに)、更新され、Bモードのとき(即ち符号器501-2からの出力が選択されるとき)は、更新されずにその値を保持するようになっている。

その後は、この品質評価・符号器選択部502による選択結果に基づき、セレクト517が再生音声品質の最適な符号器501-i( $i=1$ 又は2)の出力を選択して、この出力を伝送路503へ伝送する。

そして、復号器518では、伝送路503を介して送られてきた符号化出力(符号器501-1からの残差信号情報と予測パラメータあるいは符号器501-2からの残差信号情報)と符号器番号情報iとに基づき再生信号を出力するのである。

以上の動作において、受信側に送出する情報としては、

由による。すなわち、Aモードが選択される場合は、線形予測分析部506で常に現在のフレームに対応する予測パラメータが計算される。従って、良好なスペクトル特性が得られるため、時間域での歪を示すセグメンタル $S/N$ が良いという条件のみで足りる。これに対して、Bモードが選択される場合は、たとえ時間域での歪を示すセグメンタル $S/N$ が良かったとしても、それはたまたまBモードでの再生信号量子化利得がよっただけによる場合があり、その場合、係数メモリ513からの予測パラメータで定まる現フレームのスペクトル特性が線形予測分析部506からの予測パラメータで定まる現フレームの真のスペクトル特性から大きくはずれている場合があり得る。すなわち、Bモードの時間域での歪は、Aモードの時間域での歪より少なくなっているにもかかわらず、係数メモリ513から得られる予測パラメータは前のフレームに対応しているものであるから、現在のフレームの予測パラメータは前のフレームのものとは全く異なることになる。このような場合には、復号側で再生される再生信号は、聴感上大きなスペクトル歪を有してしまう。従って、Bモードが選択される場合には、時間域での歪だけでなく、周波数域での歪、すなわちLPCケブストラム距離CDに基づくスペクトル歪についても評価するのである。

符号器501-2のセグメンタル $S/N$ が符号器501-1のセグメンタル $S/N$ より良ければ、現在のフレームのスペクトル特性は前のフレームのスペクトル特性とあま

Aモードでは残差信号情報であるコード番号とベクトル量子化した予測パラメータ(LPCパラメータ)等であり、Bモードではコード番号等でLPCパラメータは送らない。但しトータルのビット数は同じとする。ここで、コード番号とは、白色雑音コードブック507-1又は507-2において、どの残差波形パターンすなわちコードベクトルを選択したかを示す情報である。そして、符号器501-1内の白色雑音コードブック507-1は収容している残差波形パターン(コードベクトル)の数が少数で、コード番号のビット数は少なく、これに対して符号器501-2内の白色雑音コードブック507-2の収容コード数は多数でコード番号のビット数も多数とする。従って、Bモードの方が入力信号と近似度が高い信号サンプルがある確立が高い。

全体の伝送ビットレートが4.8Kbpsの場合を例にして、前述の第2図の第2の従来例と第5図の第2の実施例において、1フレームあたりの伝送ビット割り当ての例を示すと第7A図及び第7B図となる。

第7A図及び第7B図から分かるように、第7B図の本実施例でもAモードでは各情報に割り当てられるビットレートは第7A図の第2の従来例と殆ど同じであるが、BモードではLPCパラメータを伝送しないのでその分をコード番号やゲイン番号に廻すことができ、これにより再生音声の品質の改善が期待できる。

以上説明したように本実施例では、音声の予測パラメー

タに余り変化がないフレームでは該パラメータを伝送せず、これにより生じた余裕分を音質向上のためのデータ送出（残差信号を多ビットにする、線形コード変を大容量にしてコード番号のビット数を大にする等）に振り分けるので、受信側再生音声信号の品質改善が図れる。

換言すれば、この実施例では、人間の発音機構における音源部と声道部との動特性に応じて、符号器の音源情報と声道情報との伝送比率を制御することができるため、低い伝送速度においても、S/Nが劣化することがなく、これにより良好な音声品質を維持できるものである。

なお、符号器501-1、501-2については、共に残差信号情報と予測パラメータとを出力するものでもよく、この場合は両符号器での残差信号情報と予測パラメータとの各割当ビット比を異なったものにする。

もちろん、符号器を3個以上設け、これらの符号器について、残差信号情報と予測パラメータ情報とを共に出力するもの（但し両情報の割当ビット比は符号器ごとに異なったものにする）や残差信号情報のみを出力するものを混在させるようにしてもよい。また、符号器における再生音声品質評価のために、再生音声信号の波形歪及びスペクトル歪を併用する場合のほか、再生音声信号の波形歪だけを用いたり、再生音声信号のスペクトル歪だけを用いたりしてもよい。

（産業上の利用可能性）

以上詳細に説明したように、本発明のモード切替型音声符号化装置によれば、音源情報と声道情報との間の伝送比率について複数種類のモードを用意し、これらの間で最適な再生音声品質が得られるモードに切り替えることが行われるので、人間の発音機構における音源部と声道部との動特性に応じて、符号器の音源情報と声道情報との伝送比率を制御することができ、これにより低い伝送速度においても良好な音声品質を維持できる利点がある。

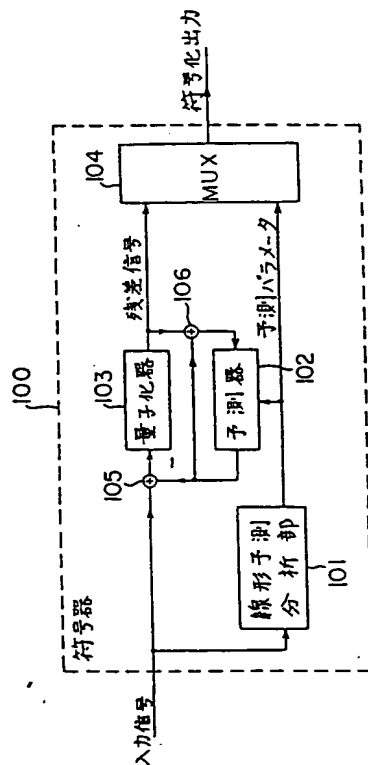


Fig. 1

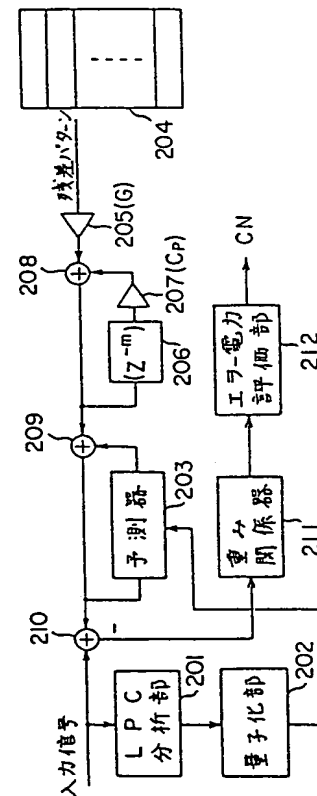


Fig. 2

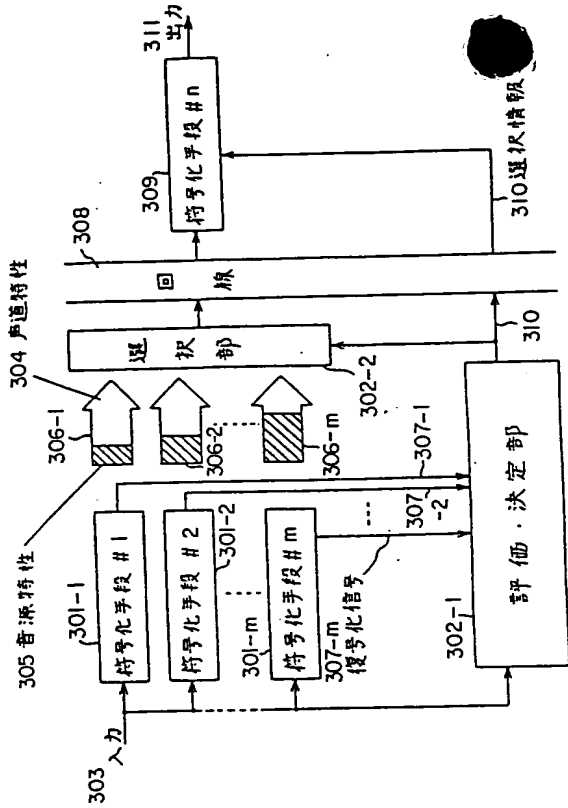


Fig. 3

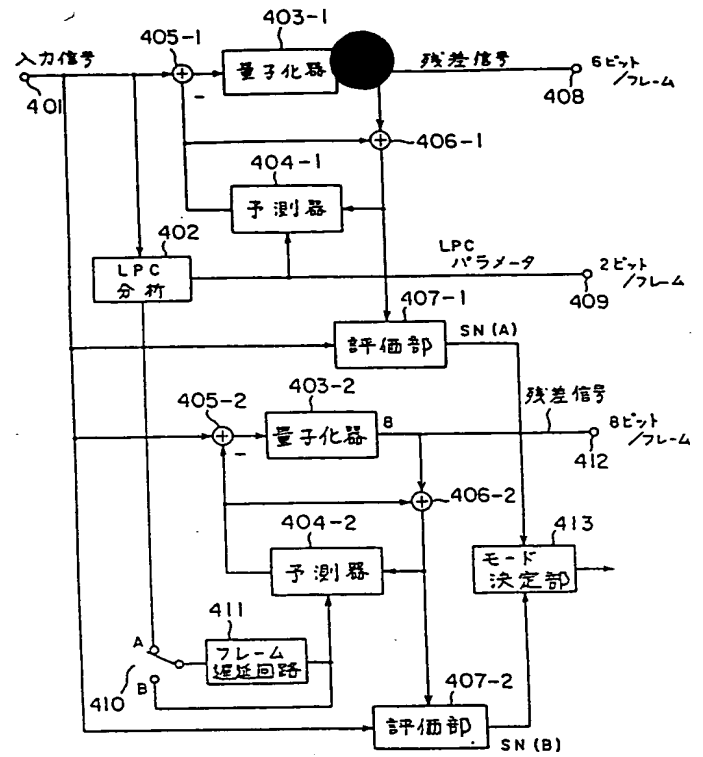


Fig. 4

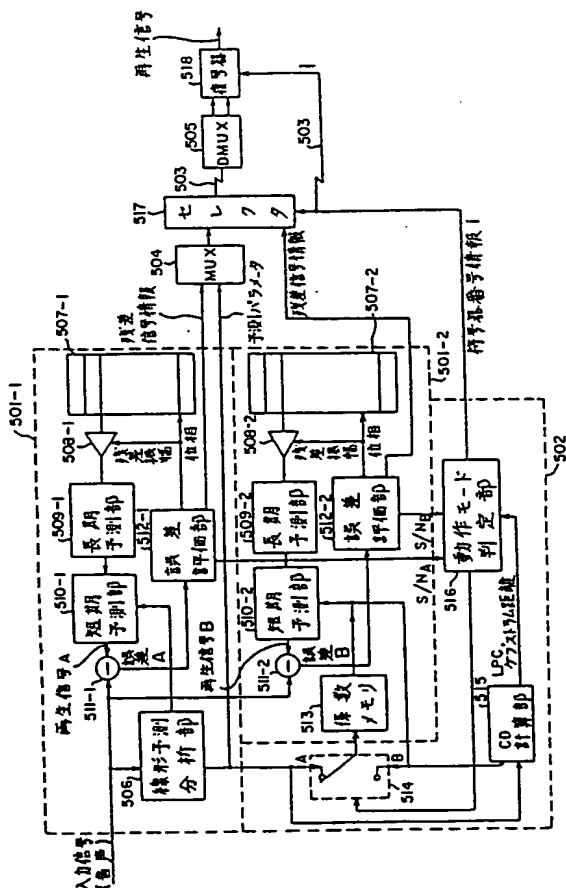


Fig. 5

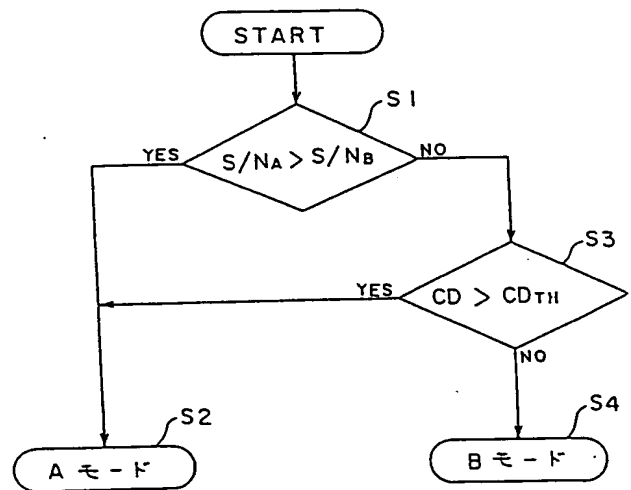


Fig. 6

コード番号	1600 bps
ゲイン	1000
ピッチ周波数	600
ピッチ係数	600
LPCパラメータ	1000
合計	4800 bps

Fig. 7A

	Aモード	Bモード
コード番号	1600 bps	2200 bps
ゲイン	1000	1350
ピッチ周波数	600	600
ピッチ係数	600	600
LPCパラメータ	950	—
モード信号	50	50
合計	4800bps	4800 bps

Fig. 7B

1. CLASSIFICATION OF SUBJECT MATTER		
IPC Class. G 10 L 9/24		
2. FIELD SEARCHED		
IPC Class. G 10 L 9/14		
3. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Relevant to Claim No. 1	Relevant to Claim No. 2
A	IEEE Transactions on Acoustics, Speech, and Signal Processing, volume ASSP-31, no. 3, June 1983, IEEE, (New York, US); P.E. Papamichalis et al.: "Variable rate speech compression by encoding subsets of the PARCOR coefficients", pages 706-712 see Abstract	1,2,4-9
A	ICC'84, Links for the Future, Science, Systems & Services for Communications, IEEE International Conference on Communications, 14-17 May 1984, Amsterdam, NL, Proceedings, volume 3, IEEE/Elsevier Science Publishers B.V. (North-Holland), (New York, US); L.B. Almeida et al.: "Harmonic coding: an introduction", pages 1169-1173 see page 1170, right-hand column, lines 23-49	1,6,7,9
<p>* Special categories of cited documents: "</p> <p>"A" documents defining the general state of the art which is not considered to be of particular importance</p> <p>"B" documents published on or after the international filing date</p> <p>"C" documents which may have priority claims or which are cited to establish the background art or to establish the state of the art at the time of the invention</p> <p>"D" documents appearing to be prior art documents, but, according to the examiner, not being prior art documents</p> <p>"E" documents published prior to the international filing date but later than the priority date</p> <p>"F" documents published after the international filing date or priority date and not in conflict with the invention but also to underline the principle or theory underlying the invention</p> <p>"G" documents of particular relevance: the claimed invention seems to be concerned by one or more of them</p> <p>"H" documents of particular relevance: the claimed invention seems to be concerned by one or more of them</p> <p>"I" documents of particular relevance: the claimed invention seems to be concerned by one or more of them</p> <p>"J" documents of particular relevance: the claimed invention seems to be concerned by one or more of them</p>		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search		Date of Mailing of the International Search Report
5th October 1989		02 NOV.1989
International Searching Authority		Signature of Authorizing Officer
EUROPEAN PATENT OFFICE		T.K. WILDS

Form PCT/ISA/210 (Revised January 1988)

A. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category	Relevant to Claim No. 1	Relevant to Claim No. 2
A	ICASSP 82, Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 3-5 May 1982, Paris, FR, volume 1 of 3, IEEE, (New York, US); A. Lacroix et al.: "A vocoder scheme for very low bit rates (quality evaluation)", pages 618-621 see pages 619-621: "Spectral distance measures and resonance locations"	2,4,5

Form PCT/ISA/210 (Revised January 1988)

第1頁の続き

優先権主張 ⑨平1(1989)3月14日⑩日本(JP)⑪特願 平1-61533

⑫発明者 田中 良紀  
⑫発明者 大田 恭士

神奈川県横浜市戸塚区汲沢8-28-1

神奈川県川崎市中原区下小田中1674番地 第八中原寮